

Prescriptive analysis of Delhi air quality

1)Introduction:

From,past two years Delhi is facing huge problem of poor air quality.Delhi air quality is so poor that they required to open fresh air cafe in their city.As Data Scientist we took the data from 2015 to find out the source of this pollution.We took Time series data of Delhi air quality to analyse which harmful substance is concentrated in Delhi's air quality.

2) Methodology:

Here,we followed standard methodology for data analysis which is multivariate analysis technique as our data is multivariate.

Here is the sample of dataset which we have taken for analysis.

	Stn Code	Sampling Date	State	City/Town/Village/Area	Location of Monitoring Station	Agency	Type of Location	SO2	NO2	RSPM/PM10	PM 2.5
0	55	05-01-15	Delhi	Delhi	Nizamuddin, Delhi	Central Pollution Control Board	Residential, Rural and other Areas	4.0	44.0	203.0	NaN
1	55	08-01-15	Delhi	Delhi	Nizamuddin, Delhi	Central Pollution Control Board	Residential, Rural and other Areas	4.0	45.0	214.0	NaN
2	55	13-01-15	Delhi	Delhi	Nizamuddin, Delhi	Central Pollution Control Board	Residential, Rural and other Areas	4.0	47.0	182.0	NaN
3	55	16-01-15	Delhi	Delhi	Nizamuddin, Delhi	Central Pollution Control Board	Residential, Rural and other Areas	4.0	43.0	204.0	78.0
4	55	21-01-15	Delhi	Delhi	Nizamuddin, Delhi	Central Pollution Control Board	Residential, Rural and other Areas	4.0	39.0	192.0	83.0

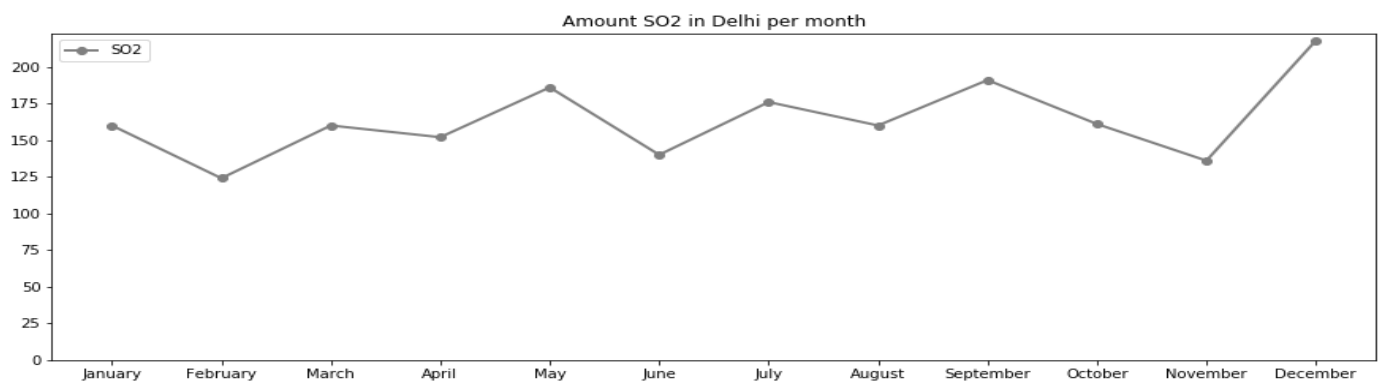
Feature explanation:

- 1)Stn_code-This column represents station code of state.
- 2)Sampling_Date-This column represents Sampling Date on which the observation was recorded.
- 3)State-This column represents State on which data collection was done.
- 4)City-This column represents City of State.
- 5)Location-This column represents Area in the City where the data was collected.
- 6)Agency-This column represents Agency name who collected the data
- 7)Type_of_Location-This column represents whether location is residential or industrial.
- 8)SO2-This column represents amount of SO2 on particular day.
- 9)NO2-This column represents amount of NO2 on particular day.
- 10)RSPM-This column represents amount of RSPM on particular day.

11)PM2.5-This column represents amount of PM2.5 on particular day.

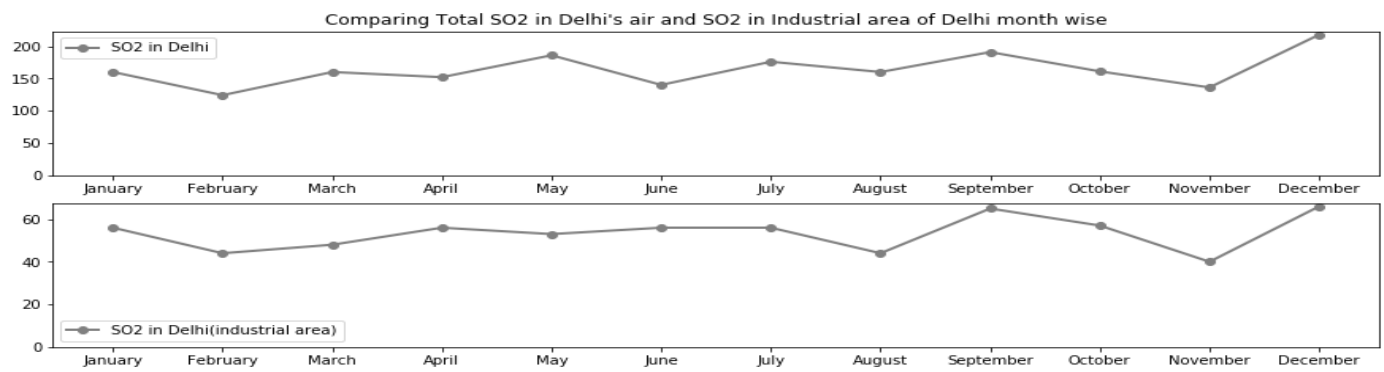
Time based analysis:

Q1)What is the concentration of SO2 in Delhi's Air?



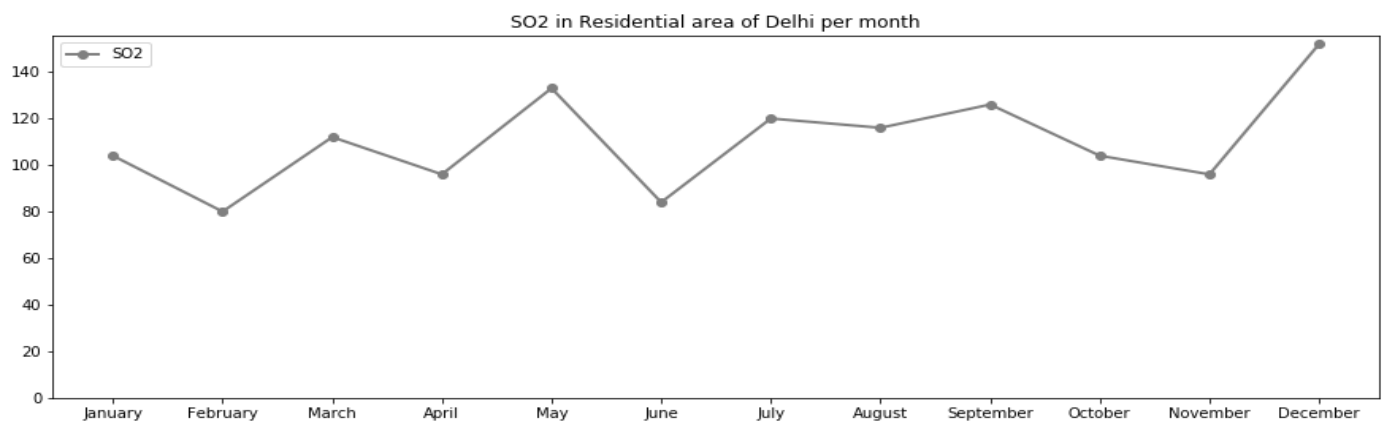
Inference:- It is observed that the amount of SO2 was lowest in November inspite of Diwali festival due to the awareness and ban done on it by the government.

Q2)Is there any relation between Total SO2 in Delhi and SO2 generated by Industrial area of Delhi?



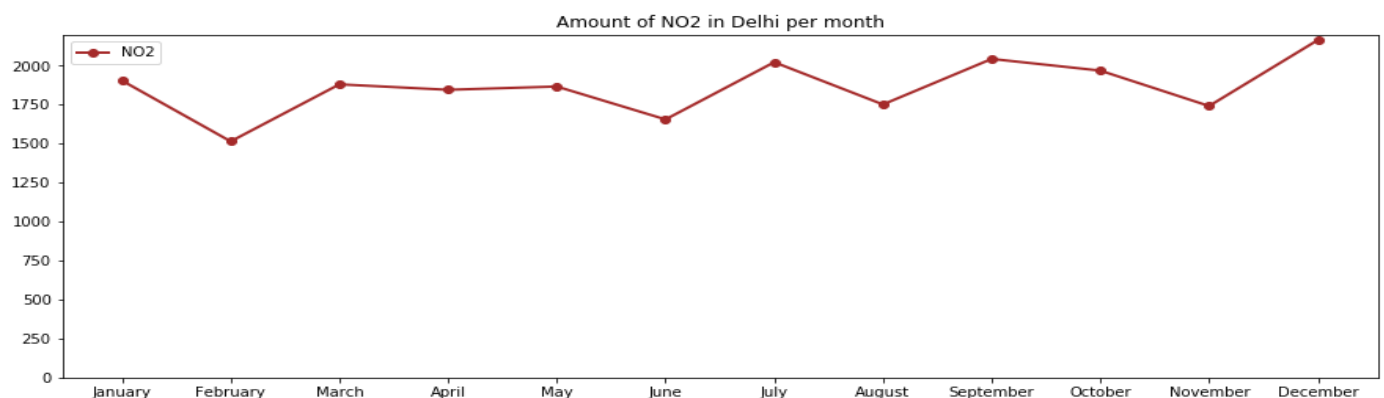
Inference: So we can see that there is a strong relation between total SO₂ and industrial area in Delhi. So, we can conclude that the main area due to pollution is industrial area or we can say that the main cause of pollution is industrial area in Delhi.

Q3) What is the concentration of SO₂ in residential area of Delhi's air?



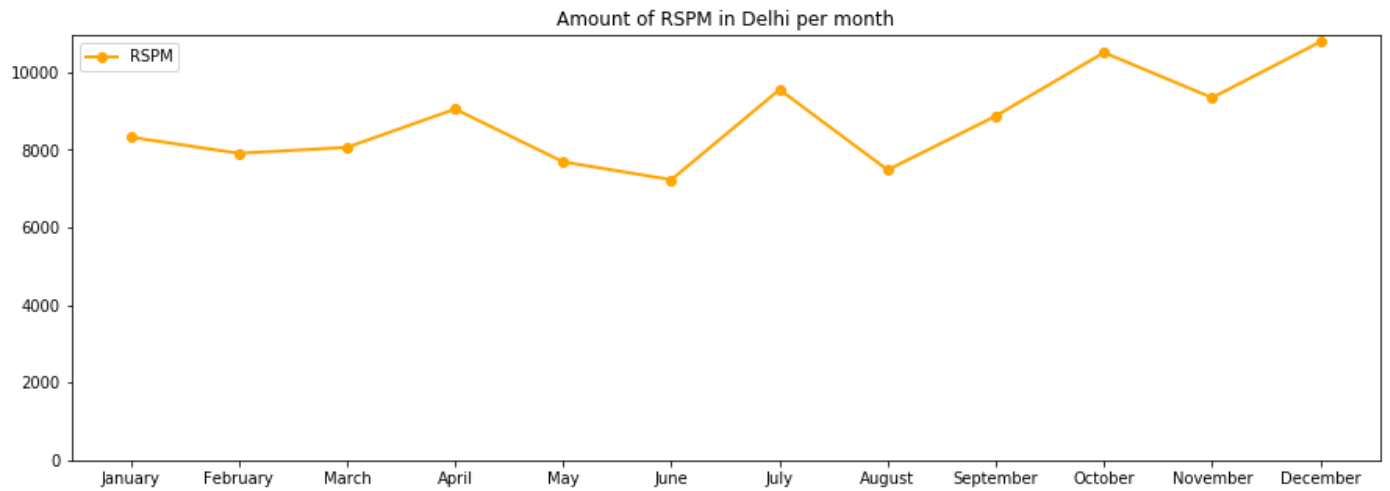
Inference: So, here we can see that the pollution in Delhi was constantly fluctuating during every month in Residential Area.

Q4) What is the concentration of NO₂ in Delhi's air?



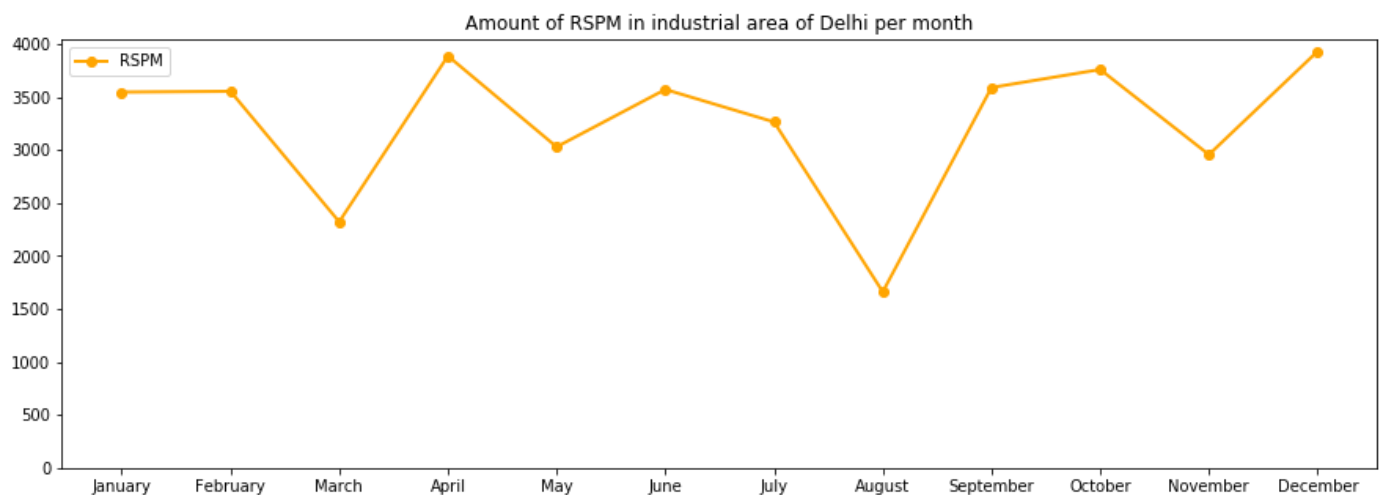
Inference: So, we can observe that the concentration of NO₂ was low during the months of February, June, August, and November.

Q5)What is the concentration of RSPM in Delhi's air?



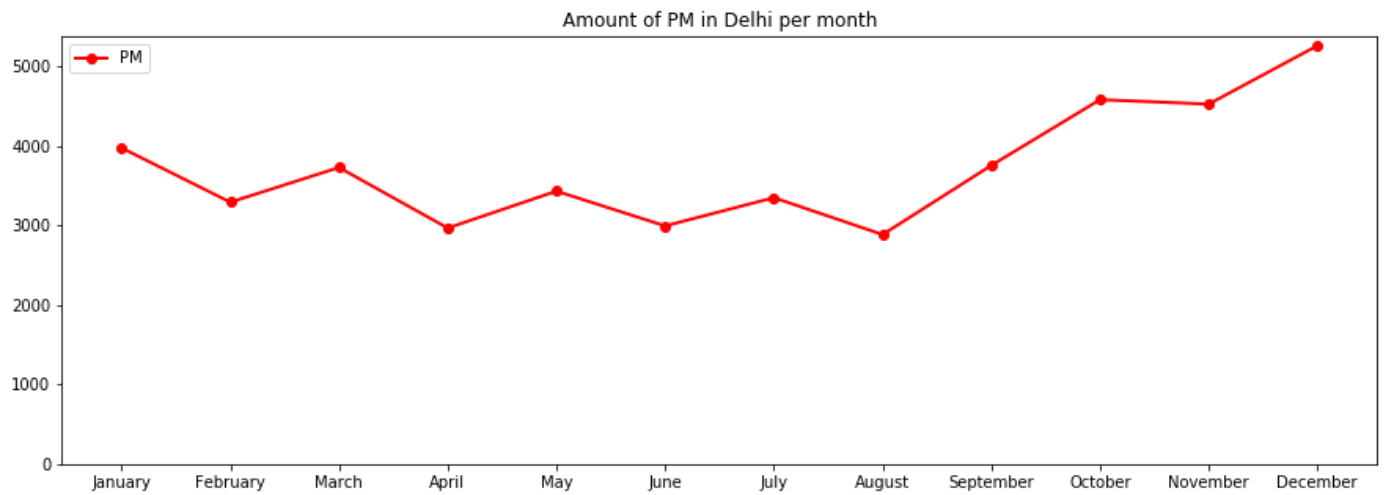
Inference:

Q6)What is the concentration of RSPM in industrial area of Delhi's air?



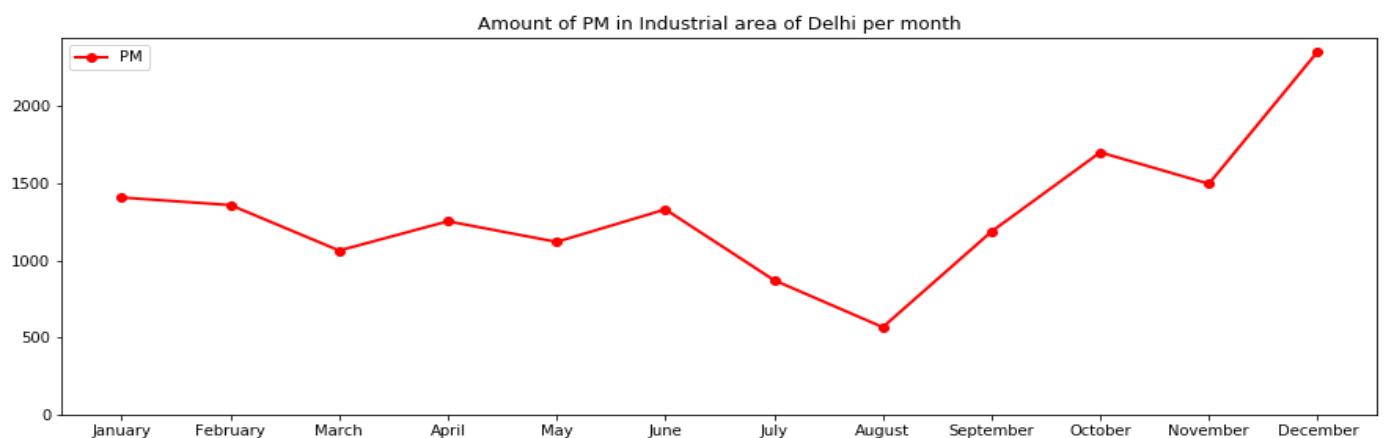
Inference:So we can observe that RSPM was least during the month of August.

Q7)What is the concentration of PM in Delhi's air?



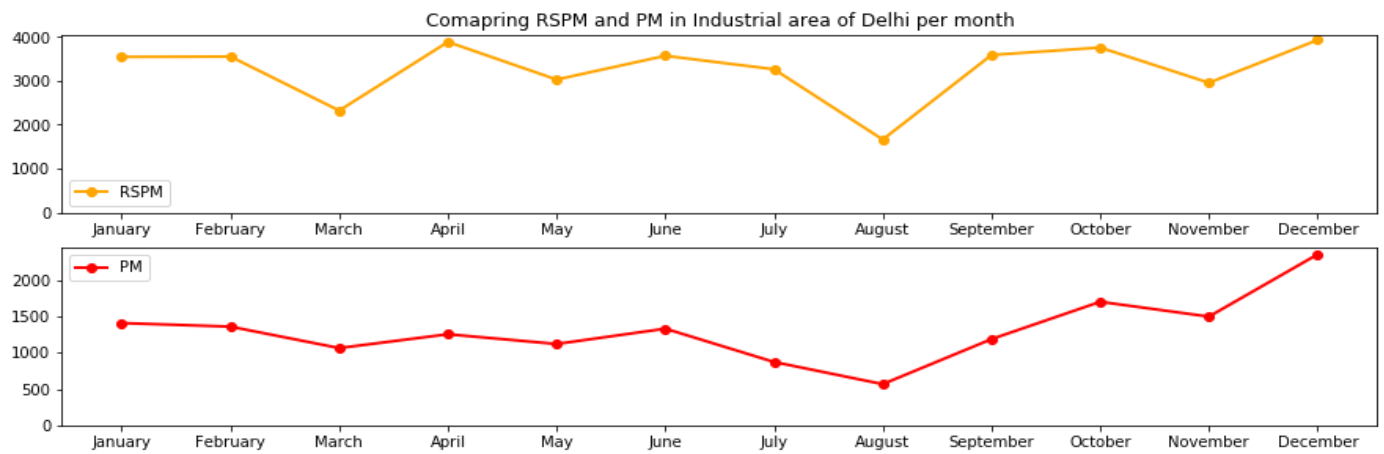
Inference: So, here we can see that the amount of PM was between 3000 to 4000 till the month of August but it suddenly rises to 4800 in the month of November.

Q8) What is the concentration of PM in Industrial area of Delhi's air?



Inference: It was observed that PM was least during the month of August due to.

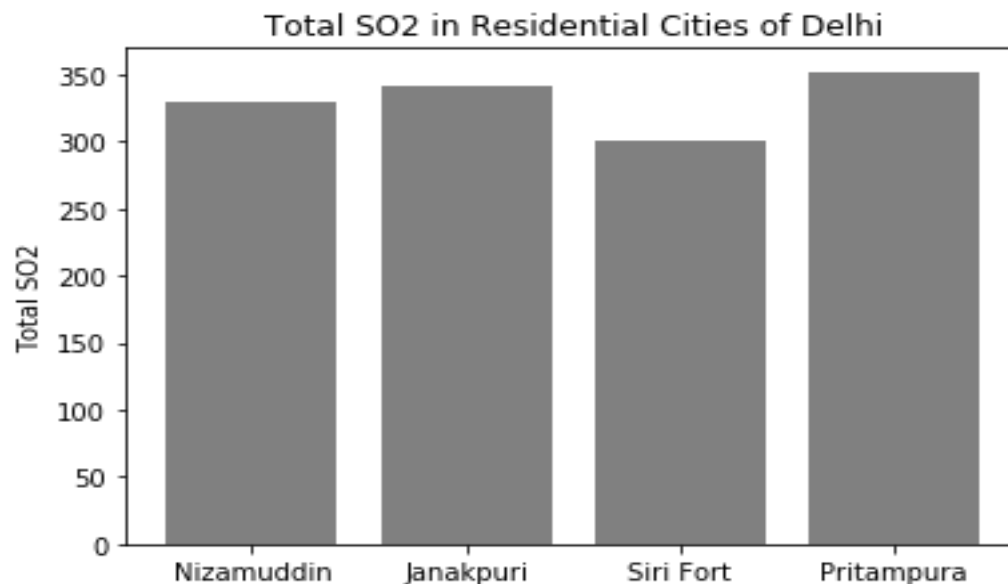
Q9) What is the reason of drastic reduction in PM and RSPM in Industrial area of Delhi in the month of August?



Inference: As, we can see that there was a drastic decrease of PM and RSPM during the month of August due to

Area based analysis

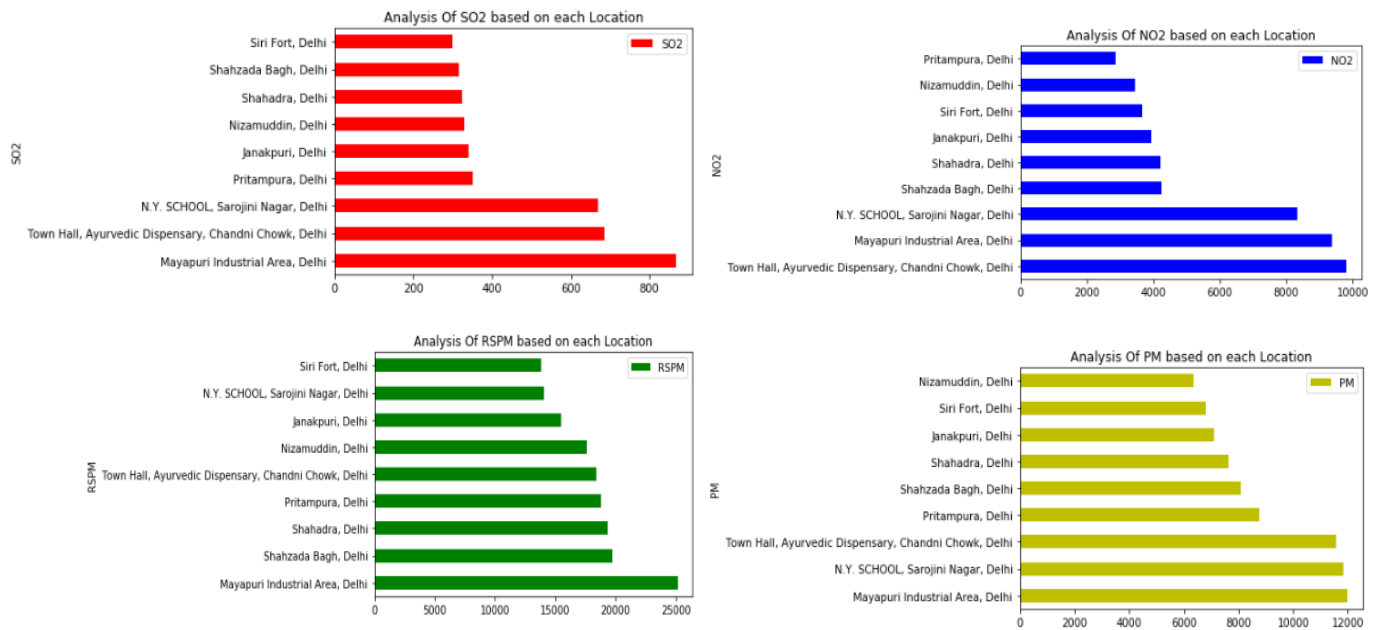
Q10) Which Residential Area is highly polluted in Delhi?



Inference: So, in the graph we can observe that concentration of SO2 was highest in Pritampura.

Location based analysis

Q11) Which location has the highest amount of pollution ?



Inference- So, here we can observe that the amount of SO₂, RSPM and PM were highest in Mayapuri Industrial Area but the amount of NO₂ was low because it is POSH area.